Stochastic Processes and their Applications 37 (1991) 331-338
North-Holland

Pascal processes and their characterization

F.T. Bruss

Department of Mathematics, University of California, Los Angeles, CA 90024-1555, USA

L.C.G. Rogers

Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, CB2 1SB Cambridge, UK

Received 4 December 1989
Revised 26 June 1990

Let (Π_t) be a counting process on $\mathbb{R}^+$ with the property that for any t, T with $0 \leqslant t \leqslant T$ the distribution of Π_T given the past $\mathscr{F}_t$ is Pascal (negative binomial) with one parameter being $\Pi_t + 1$ and the probability parameter depending only on t and T. Does such a process exist? If so, how is it characterized? Finally, what is the most convenient way to model such a process? These questions are motivated by the distinguished role of the Pascal distribution in finding explicit solutions of optimal selection problems based on relative ranks. We answer them completely.

Yule processes * mixed Poisson processes * record processes * martingales

1. Introduction

We consider in this paper three counting processes:

(i) The Yule process, defined to be the continuous time Markov chain on $\mathbb{N} = \{1, 2, \ldots\}$ whose Q-matrix has non-zero entries

$$q_{i,i+1} = -q_{i,i} = i, \quad i \in \mathbb{N}.$$

Let $(Y_t)_{t \geqslant 0}$ be such a process with $Y_0 = 1$.

(ii) The Poisson process with randomized rate: if $(N_t)_{t \geqslant 0}$ is a homogeneous Poisson process with rate 1, and V is an independent negative exponential random variable of mean 1, we let

$$\tilde{N}_t = N_{V(e^t - 1)}$$

denote the Poisson process with rate randomized by V.

Research supported in part by the Commission of the European Communities (Grant No. ST2J-0277-C) and the Alexander von Humboldt Foundation.

(iii) The Pascal process defined as follows.

Let $(\Pi_t)_{t\geq 0}$ be a counting process with the distributional prescription that for all $T>0$ and all $0\leq t\leq T$,

$$P(\Pi_T=n\,|\,\mathcal{F}_t)=\binom{n}{\Pi_t}(\xi_t(T))^{\Pi_t+1}(1-\xi_t(T))^{n-\Pi_t} \tag{1.1}$$

where $\mathcal{F}_t=\sigma(\{\Pi_s:s\leq t\})$, $\Pi_0=0$ and where $0<\xi_t(T)<1$ for all $0<t<T$ and $\xi_t(t)=1$ for all $t>0$.

The first two processes are more familiar than the third, so let us briefly indicate the reason for interest in the latter, which arises naturally in the context of optimal selection strategies based on relative ranks, when the number N of options is unknown. In particular, the example we are about to discuss proves that Pascal processes exist.

Let $X_1, X_2, \ldots, X_N$ be continuous i.i.d. random variables with values in $[0, T]$ and common c.d.f. F, where N is an $\mathbb{N}$-valued random variable independent of the X_i's. Let $N(T)=N$ and

$$N(t)=\#\{X_i: X_i\leq t\},\quad 0\leq t\leq T,$$

and focus our interest on the posterior distribution

$$P(N=n\,|\,\mathcal{F}_t)$$

where $\mathcal{F}_t$ denotes the σ-algebra generated by $\{N(s): s\leq t\}$. This posterior distribution depends clearly on the prior $\{P(N=n)\}_{n=0,1,\ldots}$ and on F, but the 'history-dependence' reduces itself to the parameter t and the observation $N(t)$ (because of the i.i.d. assumption). A straightforward application of Bayes formulae yields then

$$P(N=n\,|\,\mathcal{F}_t)=\frac{\binom{n}{N(t)}(F(t))^{N(t)}(1-F(t))^{n-N(t)}P(N=n)}{\sum_{k=N(t)}^{\infty}\binom{k}{N(t)}(F(t))^{N(t)}(1-F(t))^{k-N(t)}P(N=k)}. \tag{1.2}$$

If we choose a geometric prior, $P(N=n)=q^n p$ say, it is easy to check that the right-hand side of (1.2) becomes

$$\binom{n}{N(t)}(1-(1-F(t))q)^{N(t)+1}(q-qF(t))^{n-N(t)}, \tag{1.3}$$

i.e. the described Pascal distribution with parameters $N(t)+1$ and $\xi_t(T)=(1-(1-F(t))q)=p-qF(t)$. Furthermore one can show that there is only one other prior producing (in this model) a Pascal posterior, namely the improper prior $P(N=n)=1$ for all n (Bruss and Samuels, 1990, Section 2.4). So much for an example of a specific model, where Pascal posteriors arise, but what are the processes with this distributional prescription? The answer is important in the field of optimal selection strategies based on relative ranks, where Pascal posteriors for an unknown number of options N present the only case where the optimal strategy is, for all loss functions, quasi-stationary, which means that the expected loss by stopping at time t on some relative rank r depends on $\mathcal{F}_t$ only through r and t.

2. Results

Our main results of this paper answer this question, namely:

Theorem 1. *The processes* $(Y_t - 1)_{t\geq 0}$, $(\tilde{N}_t)_{t\geq 0}$ *and the Pascal process with* $\xi_t(T) = \exp(-(T-t))$ *are the same.*

The Pascal process with $\xi_t(T) = \exp(-(T-t))$ will be called the *standard* Pascal process. The next result completes the characterization, namely

Lemma 1. *Every Pascal process can be reduced to a standard Pascal process by a deterministic time change, and every strictly-increasing continuous deterministic time change of a standard Pascal process is again a Pascal process.*

The equivalence of $(Y_t - 1)_{t\geq 0}$ and $(\tilde{N}_t)_{t\geq 0}$ seems also to be new, and is of interest on its own; the helpful comments of an anonymous referee led us to this result. The fact that the Markov chain $Y_t - 1$ can be modelled by a homogeneous Poisson process, a deterministic time change and a *single* randomization is indeed remarkable. We shall prove this by showing that the martingale representations of both processes are equivalent.

Our final results deal with those distinguished properties of Pascal processes which are, as we mentioned, very useful in optimal choice problems. Let $X_1, X_2, \ldots$ be a sequence of continuous i.i.d. random variables and consider temporarily the bivariate process $(T_k, X_k)_{k=1,2,\ldots}$, where T_k denotes the kth occurrence time of a Pascal process. Call T_k a j-record time, if X_k is the jth largest among $X_1, X_2, \ldots, X_k$ (see e.g. Nevzorov, 1988; for preliminary work see Rényi, 1962). Here $1 \leq j \leq k$. Collect all j-record times in a point process $(R_t^j)_{t\geq 0}$. Then we have the following surprising result.

Theorem 2. *The record-times processes* $(R_t^j)_{t\geq 0}$, $1 \leq j \leq k+1$, *of a Pascal process are i.i.d. on* $(T_k, \infty]$.

Our proof will be based on the equivalence of the standard Pascal process and the mixed Poisson process with randomized $V \sim \exp(1)$, and on Ignatov's theorem.

Finally, we will prove the following extension of Theorem 2.

Theorem 3. *The record-times processes* $(R_t^j)_{t\geq 0}$, $1 \leq j \leq k+1$, *of a Pascal process are, on* (T_k, ∞), *i.i.d. inhomogeneous Poisson processes with intensity function* $\lambda(t)$ *satisfying*

$$\int_t^T \lambda(s)\,\mathrm{d}s = -\ln(\xi_t(T)).$$

3. Proofs

Proof of Theorem 1. Let $(Y_t)_{t\geqslant 0}$ be a Yule process with initial state $Y_0 = 1$ and parameter $\beta = 1$, as defined in Karlin and Taylor (1975, p. 123). Set $\Pi_t = Y_t - 1$, then by (1.7) (p. 123) Π is a standard Pascal process.

In order to show that $(Y_t - 1)_{t\geqslant 0}$ and $(\tilde{N}_t)_{t\geqslant 0}$ as defined in (ii) are equivalent, we use the basic equation of filtering (see e.g. Rogers and Williams, 1987, Section VI8). Let $\mathscr{F}_t$ be the filtration generated by $\tilde{N}_t$ and V, and let $\mathscr{G}_t$ be the filtration generated by $\tilde{N}$ alone. Then

$$\tilde{N}_t - \int_0^t e^s V \, ds \quad \text{is a } (\mathscr{F}_t)\text{-martingale} \tag{3.1}$$

and thus

$$\tilde{N}_t - \int_0^t E(e^s V \mid \mathscr{G}_s) \, ds \quad \text{is a } (\mathscr{G}_t)\text{-martingale.} \tag{3.2}$$

The nice interpretation is that we see $\tilde{N}$ and are trying to make inferences about the unknown parameter V.

Now

$$P(V \in dy \mid \tilde{N}_t = k) = \frac{P(\tilde{N}_t = k \mid V \in dy) \cdot P(V \in dy)}{\int_{\mathbb{R}^+} P(\tilde{N}_t = k \mid V) \, dP(V)} \tag{3.3}$$

where the conditional law of $\tilde{N}_t$ given $V = y$ is Poisson with parameter

$$\mu(t) = \int_0^t e^s y \, ds = y(e^t - 1). \tag{3.4}$$

Using (3.4) in (3.3) and substituting $w = v\, e^t$ in the integral on the right-hand side of (3.3) shows quickly that

$$P(V \in dy \mid \tilde{N}_t = k) = (y^k \, e^{-y \exp(t)} \, e^{t(k+1)}/k!) \, dy$$

from which we obtain similarly, by integration

$$E(V \mid \mathscr{G}_t) = (\tilde{N}_t + 1) \, e^{-t}.$$

Thus, by equation (3.2),

$$\tilde{N}_t - \int_0^t (1 + \tilde{N}_s) \, ds \quad \text{is a } (\mathscr{G}_t)\text{-martingale}$$

which is to say that

$$(\tilde{N}_t)_{t\geqslant 0} \quad \text{is a Yule process,}$$

and the proof is complete. □

Proof of Lemma 1. The right-hand side of (1.1) is the probability that the $(\Pi_t + 1)$st success in a sequence of independent Bernoulli trials with success probability $\xi_t(T)$

each is obtained at the $(n+1)$st trial, so that Π_T counts the number of trials *before* this last success, in other words

$$E(\Pi_T \mid \mathcal{F}_t) = \frac{\Pi_t + 1}{\xi_t(T)} - 1. \tag{3.5}$$

Now define $Z_t \equiv \Pi_t + 1$, and use the trivial identity

$$E(Z_T \mid \mathcal{F}_t) = E(E(Z_T \mid \mathcal{F}_u) \mid \mathcal{F}_t), \quad 0 \leq t \leq u \leq T,$$

together with (3.5) which yields

$$E(Z_T \mid \mathcal{F}_t) = Z_t / \xi_t(T) = Z_t / (\xi_u(T)\xi_t(u)). \tag{3.6}$$

These equalities imply that for all $0 \leq t \leq u \leq T$,

$$\xi_t(T) = \xi_t(u) \cdot \xi_u(T) \tag{3.7}$$

or equivalently that there exists some increasing function $h : \mathbb{R}^+ \to \mathbb{R}^+$ with

$$\xi_t(T) = \exp\{h(t) - h(T)\}. \tag{3.8}$$

It follows from (3.5) that $\lim_{T \downarrow t} \xi_t(T) = \xi_t(t) = 1$ for all t, so that h is right continuous. Moreover, since Π is a counting process, P (for some t, $\Delta\Pi_t > 1) = 0$, which implies that $\lim_{t \uparrow T} \xi_t(T) = 1$, and h is left continuous, whence h is continuous. Consequently h^{-1} is well-defined, $h(h^{-1}(t)) = t$, and the first statement of Lemma 1 is proved. The second follows easily. □

Proof of Theorem 2. To be quite explicit about the statement of Theorem 2, we shall prove that for each $t > 0$, conditional on $\{T_k < t\}$, the restrictions to (t, ∞) of R^j, $1 \leq j \leq k+1$, are i.i.d. Let $X_1, X_2, \ldots$ be continuous i.i.d. random variables with a common c.d.f. F concentrated on $\mathbb{R}^+$, $F(x) < 1$ for all x. Further let $N^1, N^2, \ldots$ be the record processes of this sequence; the N^j are i.i.d. point processes by Ignatov's theorem and they are Poisson processes if F is exponential. Fixing $0 < a < b$, let $T \equiv \inf\{n \in \mathbb{N}: X_n > b\}$. Set $\tilde{N}^1, \tilde{N}^2, \ldots$ to be the point processes of records which occur at or before time T (so effectively we discard $X_{T+1}, X_{T+2}, \ldots$). Fix k and define

$$\begin{aligned} A &= \{\text{there are at least } k \text{ of the } X_i \text{ in } (a, \infty),\ 1 \leq i \leq T\} \\ &= \left\{ \sum_{r=1}^{T} I_{(a,\infty)}(X_r) \geq k \right\}. \end{aligned} \tag{3.9}$$

Notice that A is an event determined by the restrictions of $N^1, N^2, \ldots$ to (a, ∞). This is because we can reconstruct the sequence of X-values in excess of a just from $N^j|_{(a,\infty)}$-start by setting down the N^1-values in ascending order (see Figure 1), and now look at the N^2-values. From the size of the first N^2-value we can decide whereabout in the sequence of N^1-values to insert it: indeed, if y is the first N^2-value, and $x_2 < y < x_3$ we know that the 2-record value y must have occurred between the times of occurrence of the 1-record values x_3 and x_4. Proceeding this way we can reconstruct the sequence of X-values in (a, ∞).

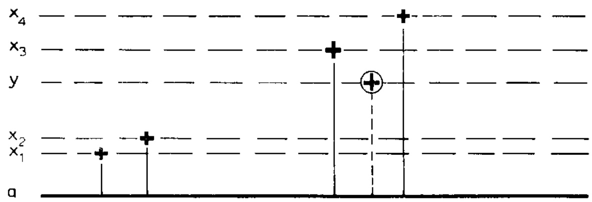

Fig. 1. N^1-values in excess of a.

Now let f be any Borel function of $\tilde{N}^1|_{[0,a]}, \ldots, \tilde{N}^k|_{[0,a]}$. Then

$$E(f(\tilde{N}^1|_{[0,a]}, \ldots, \tilde{N}^k|_{[0,a]})I_A) = E(f(N^1|_{[0,a]}, \ldots, N^k|_{[0,a]})I_A) \tag{3.10}$$

because, when event A happens, $\tilde{N}^j|_{[0,a]} = N^j|_{[0,a]}$ for $j = 1, \ldots, k$. By Ignatov's theorem, the latter equals

$$E(f(N^1|_{[0,a]}, \ldots, N^k|_{[0,a]}))P(A).$$

Thus *conditional* on A, the processes $\tilde{N}^j|_{[0,a]}$, $j = 1, \ldots, k$, are i.i.d. with the law of $N^1|_{[0,a]}$. This result, which is interesting in its own right, will accomplish the proof by recalling that a Pascal process is a Poisson process with exponentially mixed rate.

To continue the proof, we consider a planar Poisson process on $(-\varepsilon, \infty) \times (-\infty, 0)$ with expectation measure equal to Lebesgue measure. (Think of $(-\varepsilon, \infty)$ as being the time axis and $(-\infty, 0)$ as the space of qualities.) Now let V be an exponential random variable with parameter ε, say, and discard all points whose quality does not exceed $-V$. Using the well-known property of exponentially distributed spacings in Poisson processes, V has a simple geometric interpretation (see Figure 2). It is the highest-quality point to appear in the time interval $(-\varepsilon, 0)$. Reflecting in $y = -\varepsilon - x$, which leaves the probabilistic structure of the planar Poisson process unaffected, yields the following situation (see Figure 3).

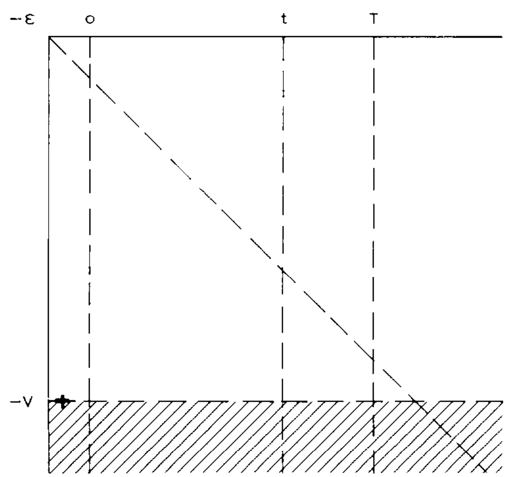

Fig. 2.

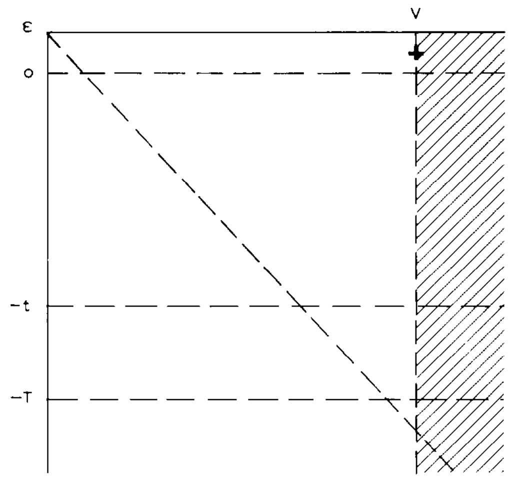

Fig. 3.

Time becomes quality and vice versa, and now V is the first time point with quality value in $(0, \varepsilon)$. Conditionally on there being at least k points in $[0, V) \times (-t, \varepsilon)$, the restrictions to $(-\infty, -t)$ of the j-record processes, $j = 1, \ldots, k+1$, are i.i.d. with the unconditional law, by the earlier argument. This completes the proof. $\square$

Proof of Theorem 3. Since all Pascal processes are obtained from a standard Pascal process by a deterministic time change, it is sufficient to prove the theorem for one Pascal process; we shall use the Pascal process which arises in the proof of Theorem 2, taking $\varepsilon = 1$.

Look again at Figure 2. The times $T_1 < T_2 < \cdots$ in $(0, \infty)$ at which there are points of the planar Poisson process in $(0, \infty) \times (-V, 0)$ are the points of a counting process N_{Vt}, where N is a standard Poisson process, and V is independent of N with an exponential distribution of mean 1. Thus if N' is a standard Pascal process, then $(N_{Vt})_{t \geq 0}$ has the same law as $(N'_{\log(1+t)})_{t \geq 0}$, whence $(N_{Vt})_{t \geq 0}$ is a Pascal process for which $\xi_t(T) = (1+t)/(1+T)$.

In the proof of Theorem 2, we saw that the record-time processes R^j, $1 \leq j \leq k+1$, restricted to (T_k, ∞) had the same law as the 1-record process R^1, so we have just got to find the law of R^1. But if we fix $0 < t < T$, and consider Figure 3, we see that the distribution of record *times* between t and T is the same as the distribution of record *values* between $-T$ and t in Figure 3. This last is easy to compute, because the values of points which fall into $(0, \infty) \times (-T, \varepsilon)$ are independent uniformly distributed in $(-T, \varepsilon)$, and are mapped by $g(x) \equiv \log(\varepsilon + T) - \log(\varepsilon - x)$ to exponential random variables of mean 1. Trivially, the record process of a sequence of independent exponentials of mean 1 is a standard Poisson process, so the number of record values in $(g(-T), g(-t))$ is a Poisson with mean $g(-t) - g(-T) = -\log((1+t)/(1+T)) = -\log \xi_t(T)$, as required. This completes the proof. $\square$

Remark. Since the three classes of processes considered in this paper are only defined in terms of their distribution, we have proved their equivalence in distribution, which is strictly all that is possible. But we have also exhibited a specific construction (on the sample space of the Poisson point process in the quarter plane), which allows all three to be realized as the same process. We are grateful to the referee for bringing this interesting observation to our attention.

Acknowledgement

Part of the research leading to this paper was done when the first author was visiting the University of Arizona as a fellow of the Alexander von Humboldt Foundation. He would like to thank the Foundation for the support and Marcel Neuts and the University of Arizona for their invitation and hospitality.

We both would like to thank the Commission of the European Communities for support through the grant 'Automatic Decision Strategies'.

References

F.T. Bruss, Invariant record processes and their applications to best choice modelling, Stochastic Process. Appl. 30 (1988) 303-316.

F.T. Bruss and S.M. Samuels, Conditions for quasi-stationarity of the Bayes rules in selection problems with an unknown number of rankable options, Ann. Probab. 18(2) (1990).

S. Karlin and H.M. Taylor, A First Course in Stochastic Processes (Academic Press, New York-San Francisco-London, 1975).

V.B. Nevzorov, Records, Theory Probab. Appl. 32(2) (1988) 211-228.

A. Reńyi, Théorie des éléments saillants d'une suite d'observations, Proc. Colloquium Comb. Methods in Probab. Theory (Aarhus Universiteit, 1962) pp. 104-115.

L.C.G. Rogers and D. Williams, Diffusions, Markov Processes, and Martingales, Vol. 2, Itô Calculus (Wiley, Chichester-New York-Brisbane, 1987).